Grenze International Journal of Engineering and Technology, January Issue

Face Presentation Attack Detection

Aditi Arora[1], Kanchan Choudhary[2] and Prabhoot Narayan Patel[3]
[1-3]Department of CSE, ABESEC, Ghaziabad, UP, India
Email: aditiarora1804@gmail.com, kanchanchoudhary0712@gmail.com, prabhootpatel@gmail.com

***Abstract*— Presentation attacks, also known as spoofing attacks, are deliberate attempts to deceive biometric systems by presenting fake inputs that bypass security mechanisms. Among these, face presentation attacks involve using fraudulent face inputs such as photos, videos, or 3D masks of legitimate users to deceive the camera sensor. This paper focuses on detecting such attacks by presenting a novel approach built on supervised learning techniques. The YOLO model was trained on 7,000 images in total, enabling it to distinguish between live and spoofed inputs with an accuracy of 94.5% when tested. The key insight leveraged by the model is the difference in gradient smoothness: the natural tone transitions in live faces are significantly different from those observed in spoofed images or videos. This technique not only enhances the detection capabilities of biometric systems but also holds potential for applications in securing online transactions by adding a layer of liveness validation during authentication processes.**

***Index Terms*— biometric systems, face spoofing, liveness detection, supervised learning, gradient analysis, online security.**

I. INTRODUCTION

Biometric systems have emerged as critical components in modern security infrastructure, providing authentication mechanisms across diverse applications ranging from smartphone access to border control. These systems identify individuals based on unique physical or behavioral characteristics such as fingerprints, facial features, iris patterns, or voice signatures. While traditional authentication methods rely on knowledge-based factors (passwords) or possession- based factors (access cards), biometric authentication offers enhanced security through inherent personal traits that cannot be easily forgotten or transferred.

However, despite their sophistication, biometric systems are not impervious to security threats. Presentation attacks, also known as spoofing attacks, represent a significant vulnerability in these systems. As illustrated in Figure 1, biometric systems are susceptible to eight primary points of attack, with the camera (sensor) representing the first and most vulnerable point in the authentication chain.

Face recognition systems are particularly susceptible to presentation attacks due to the accessibility of facial images through social media and the relative ease of creating convincing spoofing instruments. (mays alshaikhli, 2021) demonstrated that attacks targeting the camera sensor can be launched using readily available tools, requiring minimal technical expertise. These attacks typically involve presenting fraudulent facial inputs to the sensor, such as:

- Printed photographs of authorized users
- Digital images displayed on screens
- Video playback on mobile devices
- Three-dimensional masks created from various materials

Grenze ID: 01.GIJET.12.1.367

What makes these attacks particularly concerning is their external nature—they do not require in-depth knowledge of the biometric system's internal workings or sophisticated technical skills to execute. Instead, they exploit the fundamental operational principle of facial recognition systems: the comparison of captured facial data against stored templates. As (Mostaani, 2019)emphasized, addressing this vulnerability requires innovative detection methods that enhance the robustness of biometric systems while maintaining their usability and convenience.

The consequences of successful presentation attacks extend beyond unauthorized access to include identity theft, financial fraud, and compromised security in critical infrastructure. Therefore, developing effective countermeasures against face presentation attacks has become an urgent priority in biometric security research, forming the central focus of this paper.

II. MOTIVATION

In India, online banking and other authentication-reliant systems lack a consistent mechanism for verifying user liveness during login or transaction processes. This absence creates a significant loophole for fraudsters to exploit, leading to substantial monetary losses. For instance, an attacker could bypass security by presenting a spoofed face to the sensor during an authentication attempt. The motivation behind this research stems from the urgent need to integrate liveness detection as a mandatory step in online authentication processes. By developing models that effectively differentiate between live and spoofed faces, we aim to prevent presentation attacks and enhance the reliability of biometric systems. If implemented, such advancements could significantly reduce fraudulent activities, providing a more secure digital environment for individuals and businesses.

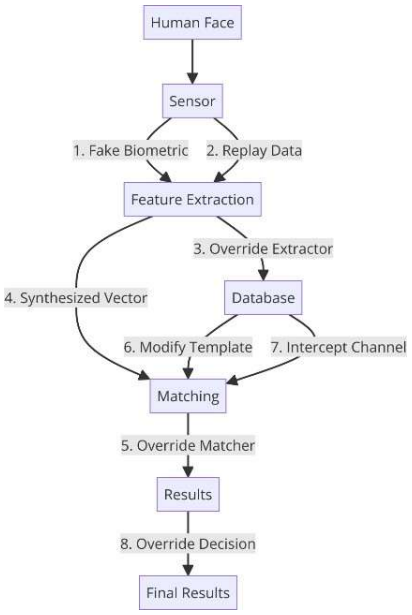

Figure 1

Figure 1. This flowchart show total of eight point of attack that are possible each attack used a different strategy to bypass the biometric system.

The main contributions of this study are summarized as follows:

- Development of a novel supervised learning approach for face presentation attack detection that achieves 94.5% accuracy on unseen data by leveraging gradient smoothness differences between live and spoofed faces
- Creation of a comprehensive dataset consisting of over 7,000 images captured under varying lighting conditions and from multiple angles to ensure robust model training across diverse scenarios
- Implementation of an efficient CNN architecture tailored specifically for detecting presentation attacks through texture analysis, edge discontinuity detection, and gradient field analysis using Sobel operators
- Deployment of a scalable system capable of processing inputs at 20 frames per second on standard GPUs, making it viable for real-time applications in biometric authentication systems

The rest of the paper is organized as follows:
Different types of instruments used is discussed in Section IV. Section V covers an overview of different approaches for face presentation attack detection. Section VI presents the detailed methodology of the proposed system, while Section VII discusses the selection of the model. Results and discussion are presented in Section VIII, followed by conclusion in Section IX. Finally, Section X outlines the future scope of this research.

III. Related Work

TABLE I

Ref	Year	Technique	Gaps	Future Scope	Advantage
[1].	2019	Multi-Channel Convolutional Neural Network (CNN) using RGB, depth, near infrared, and thermal.	Limited exploration of adversarial attacks and deepfake detection.	Incorporating adversarial training and more channels such as ultrasonic or radar for enhanced spoof detection.	Improves robustness against diverse presentation attacks by using multi-modal inputs.
[2].	2019	Deep Learning combining CNN and RNN models with depth.	Limited generalization to unseen spoofing techniques a	Extending the model to handle unknown attack patterns and training with more generalized datasets.	Effectively captures both spatial and temporal information for accurate spoof detection.
[3].	2020	Feature-based approaches for face anti-spoofing.	Struggles with handling high-quality spoofing attacks like hyper realistic masks.	Use advanced AI techniques like GANs to generate complex attack scenarios for better training models.	Enhances existing models by focusing on spoof-specific features.
[4].	2018	Deep learning, feature-based methods.	Not suited for dynamic lighting or real-time systems.	Optimizing for real-time detection with low computational costs and varying environmental conditions.	Provides better Accuracy than traditional feature-based approaches.
[5].	2021	2D and 3D presentation attacks for testing models.	Lacks a focus on mobile device applications and their spe cific challenges.	Incorporate mobile-specific attack patterns like screen reflections and mobile camera limitations.	Offers a diverse dataset for testing PAD methods, improving model performance on real-world scenarios.
[6].	2019	CNN with temporal cues for spoof detection.	Not suitable for static images or low-quality video feeds.	Extend this method to handle Lower resolution or noisy inputs, possibly via pre- processing techniques.	Effectively captures motion and liveness cues to distinguish real vs spoofed subjects.
[7].	2022	Depth map integration with deep learning to analyze 3D structure.	Depth cameras are expensive and not widely available on consumer devices.	Research cheaper alternatives to depth sensing, such as software-based depth estimation.	Improves detection accuracy by analyzing 3D face structures, making it harder to spoof with 2D images.
[8].	2023	Spoof detection model using RGB and infrared.	Infrared cameras are not Standard on most consumer devices, Limiting application scope.	Develop methods to simulate infrared data or use alternative sensors.	Works in a variety of lighting conditions, providing more robust spoof detection.

IV. Presentation Attack Instruments

Face presentation attack detection is a critical area of focus. In this section, we will examine the tools commonly used to execute such attacks. These attacks are particularly concerning because even individuals with limited technical expertise or knowledge of how biometric systems work can still pose a threat to the system. We will explore the various types of presentation attack instruments that make these systems more vulnerable.

A. Photograph or Recording Attacks

In this type of attack, the attacker uses an image or video of the target's biometric characteristic to deceive the biometric system. These attacks are relatively easy to execute and are commonly used to bypass biometric security.

Paper Face Mask
A simple face mask made of paper or other basic materials, designed to imitate the target's facial features, often used in low- tech attacks.

Resin Composite Face Mask
A more durable face mask made from resin-based materials, offering a more convincing replica of the target's face, typically used in more sophisticated attacks.

Mask with Eye Holes
A basic mask featuring cut-out eye holes, commonly used in less complex attacks. While it may look like the target's face, the lack of detail makes it easier for modern biometric systems to detect.

Hyper-realistic Face Masks
Advanced masks created using high-quality materials, such as silicone, which can closely mimic the target's facial features, posing a serious challenge to biometric systems with high precision detection capabilities.

B. Synthetic Biometric Characteristics Attacks

This type of attack involves generating synthetic models of biometric features, either based on a target's characteristics or generic templates, which are then presented to the biometric system to gain unauthorized access. These modern techniques often involve the use of advanced materials like silicone.

Generic Silicone Face Mask
A synthetic face mask created from silicone that replicates generic facial features, often designed to mimic a typical human face rather than a specific individual. These masks are commonly used to bypass facial recognition systems.

Customized Silicone Face Mask
A silicone face mask that is custom-made to replicate the unique biometric features of a specific individual. These masks are much more accurate and harder to detect by biometric systems, often used in high-stakes or high-security attacks.

C. Self-Modification Attacks

In these attacks, the individual modifies their own biometric characteristics, either physically or digitally, to replicate the target's features and deceive the biometric system. This category involves altering one's appearance or digital representation to mimic another's biometric identity.

Face Makeup
The attacker uses makeup techniques to alter their facial appearance and mimic the target's facial features. While relatively simple, this method can be effective in deceiving facial recognition systems if done skillfully.

Morphing Face Images
The attacker combines or "morphs" digital images of their own face with the target's face to create a composite image that closely resembles the target. This technique can be used to generate realistic but fraudulent facial data to bypass facial recognition systems.

Deepfake Technology
Deepfake technology involves the use of artificial intelligence and machine learning to generate hyper-realistic video or images of a person's face.
This technology can convincingly alter a person's appearance, allowing attackers to impersonate the target and evade biometric detection systems.

V. DIFFERENT APPROACH

A. Texture Analysis

Texture analysis involves examining the surface properties of the skin or facial features. It is based on the premise that real human skin has a unique texture that differs significantly from synthetic materials, such as paper, silicone, or plastic, used in fake masks or photos.

- How It Works: The biometric system analyzes the fine-grained details of the face's surface texture, such as pores, wrinkles, and skin patterns. Real skin shows complex and varied textures, while artificial materia ls typically present unnatural, repetitive patterns or overly smooth surfaces.

- Unnatural Texture: Synthetic materials may not reflect light in the same way as real skin, or may have visible seams, edges, or patterns that are inconsistent with natural skin textures.
- A photograph or mask may have noticeable pixelation, unnatural shading, or unrealistic glossiness, which can be detected using texture analysis algorithms.

Image Example:

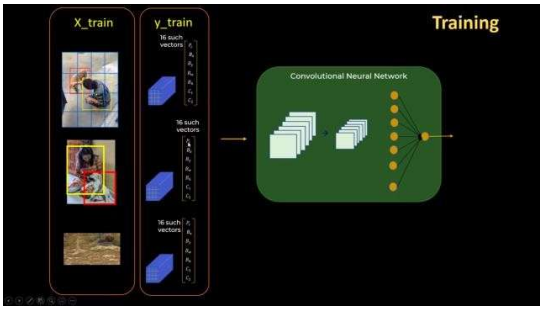

Figure 2

Figure 2. This image shows how to form input training matrix using YOLO v8.

B. Statistical Analysis

Statistical analysis helps identify inconsistencies in the color distribution and pixel patterns of a captured face. This method compares the statistical properties of the input image with known characteristics of real human faces.

- How It Works: The system performs an analysis of the image's pixel distribution and color patterns, checking for inconsistencies such as unnatural lighting, color artifacts, or odd gradient transitions.
- Inconsistencies: A real human face typically has smooth transitions between light and dark areas, whereas a fake image may have sharp boundaries, poor lighting consistency, or unnatural color shifts.
- If a face mask is presented to the system, the color distribution on the mask may differ from a normal human face, such as having a uniform color pattern or unrealistic skin tones.

C. Dynamic Analysis

Dynamic analysis involves tracking subtle natural movements, such as eye blinking or facial expressions, that are difficult to replicate with static images or fake masks.

- How It Works: The system looks for natural micro- expressions and dynamic movements that occur in real human faces, such as eye movement, blink patterns, and slight facial muscle contractions.
- Detecting Natural Micro-Expressions: Since photographs or videos of a person do not capture these micro-expressions in real-time, an attacker attempting to use a static image or video will not exhibit the same level of dynamic activity.
- Eye Movement Detection: A biometric system might ask the user to blink or follow an object with their eyes to confirm that the face is real and not just a still ima ge or a mask.

Image Example:

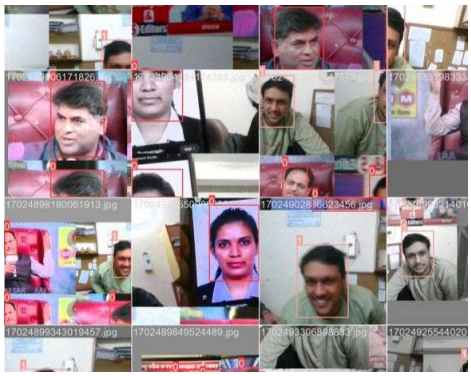

Figure 3

Figure 3. Show image which form the data set to tra in the model. In each frame there are 0 and 1 numbered in which 0 shows face and 1 show real image.

D. Motion Analysis

Motion analysis focuses on detecting movements, particularly those associated with natural human activities such as blinking, head tilts, or small facial gestures. These movements are difficult to simulate accurately in a fake biometric system.

- How It Works: The system analyzes the movement patterns in the video or image feed, looking for signs of natural movement, such as blinking, head movements, and lip movements. These are typically absent in static images or video playback of a person.
- Detecting Blinking and Eye Movements: Since real human faces exhibit regular, involuntary movements (like blinking), systems can ask users to blink or make subtle movements as part of the authentication process to verify that the face is alive.
- A video spoof of a person may show smooth, unnatural facial transitions without any blinking or head movement, triggering the motion analysis detection mechanism.

Image Example:

Figure 4. Show the results of testing in the intermediate stage.

E. 3D Depth Analysis

3D depth analysis utilizes depth-sensing technologies such as infrared sensors or structured light to capture the spatial geometry of the face. This method helps distinguish between real, three-dimensional human faces and two-dimensional photos or videos.

- How It Works: The system uses depth sensors (e.g., infrared sensors) to capture the 3D structure of the face. Real human faces have unique, detailed depth information (e.g., nose contours, cheekbones, and facial curves) that fake images or masks lack.
- Differentiating Real Faces from Fake Images: Fake images or 2D masks cannot provide accurate depth information, while real human faces have distinct depth variations that are hard to replicate.
- 3D Face Scanners can create a detailed depth map of the face, verifying the spatial dimensions and detecting fake images, which will lack true 3D features.

VI. METHODOLOGY

The proposed system is designed to address the vulnerabilities in face-based biometric systems by effectively detecting presentation attacks. The methodology is detailed below, outlining each critical step in the pipeline to ensure comprehensive implementa tion.

A. Dataset Collection and Preprocessing

The foundation of the system lies in the quality and diversity of the dataset:

- Dataset Collection: The dataset was built from scratch to include over 7,000 images categorized into live face inputs and spoofed inputs. Spoofed inputs included printed photographs, video replays on screens.

- Illumination variation: To mimic real-world scenarios, images were captured under varying lighting conditions (bright sunlight, indoor artificial light, and low-light environments) and from multiple angles.
- Gradient Analysis: Pixel-level intensity gradients were computed to identify differences in tone transitions. A gradient field was derived using Sobel operators, with live faces showing more non- uniformity compared to spoofed inputs.
- Augmentation: To increase dataset variability and robustness, transformations like flipping, random cropping, and brightness adjustments were applied.

B. Feature Extraction

The model emphasized extracting features unique to live faces:
- Texture Analysis: Fine details such as skin pores and edge irregularities were key indicators of liveness.
- Edge Discontinuities: Spoofed faces often exhibited unnatural edges due to print artifacts or screen bezels, which were detected and classified.

C. Model Architecture

The CNN architecture was tailored to process facial inputs and extract meaningful patterns:
- Input Layer: Handled grayscale images, ensuring uniformity across inputs.
- Convolutional Layers: Detected local patterns like textures, gradients, and edges. Multiple filters enhanced feature extraction for different scales and orientations.
- Pooling Layers: Reduced feature map size, retaining significant patterns while discarding unnecessary noise.
- Fully Connected Layers: Interpreted extracted features and generated output classes (live/spoofed).
- Output Layer: The final SoftMax activation ensured probabilistic predictions for each class.

D. Training and Validation

The training process prioritized accuracy and generalization:
- Loss Function: Cross-entropy loss calculated the error between predicted and actual outputs.
- Optimizer: The Adam optimizer dynamica ly adjusted learning rates, accelerating convergence.
- Validation Strategy: A hold-out validation set was used, with early stopping preventing overfitting. The final model achieved an accuracy of 94.5% on unseen data.

VII. SELECTION OF MODEL

This is Used for object detection and classification, Among all the history of models like
- Sliding window object
- R CNN
- Fast R CNN
- YOLO

YOLO v8 was preferred since it was most efficient for ima ge classification and localization.
Image classification- Image classification is a computer vision task where the goal is to assign a label or category to an input image based on its content.
Image Localization- image localization involves not only classifying the objects present in an image but also determining their locations within image.

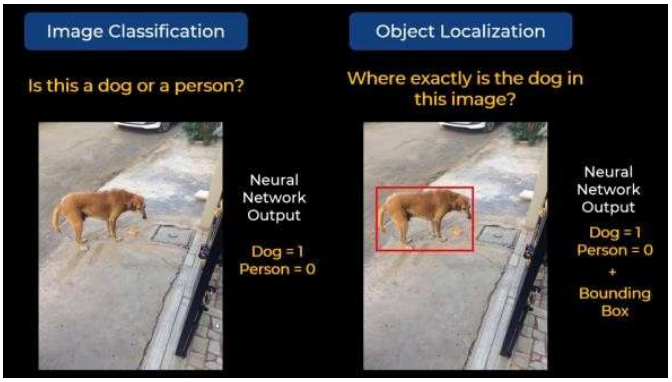

Figure 5. Show how internal working of YOLO and how ima ge classification and image localization.

VIII. RESULTS AND DISCUSSION

The model's performance was evaluated using various metrics, including accuracy, precision, recall, and F1-score. Its ability to generalize across diverse conditions was key to its success, the same is the figure 2.

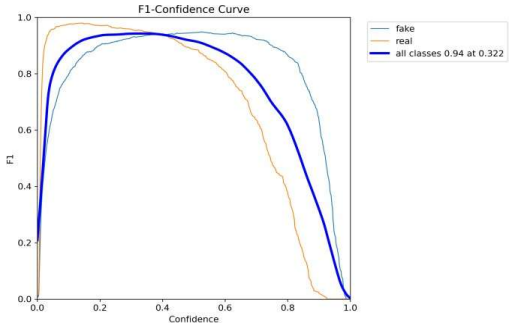

Fig 6

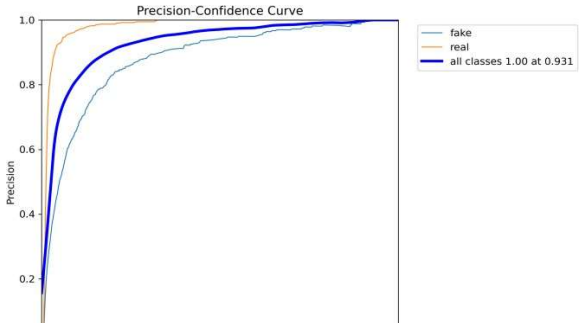

Fig 7

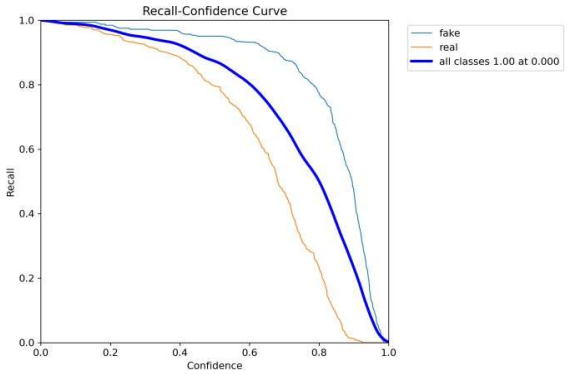

Fig 8

TABLE I

Epoch value	Patient value	Box loss	Cls loss	Dfl loss
1	10	1.0646	1.6285	1.2066
10	15	0.77124	0.47961	1.0063
117	30	0.5079	0.27986	0.9034
300	25	0.09 (expected)	0.12 (expected)	0.02 (expected)

Table 1. This table presents the final results along with the expected values of Classification Loss (CLS Loss) and Distribution Focal Loss (DFL Loss).

A. Robustness Under Different Conditions

- The model consistently performed well in varying lighting scenarios, achieving over 92% accuracy in environments ranging from bright sunlight to dimly lit interiors.
- It demonstrated resilience to minor distortions such as background noise and facial occlusions, proving effective in real-world applications.

B. Texture-Based Differentiation

Gradient analysis was instrumental in identifying spoofed inputs, also done by (mays alshaikhli, 2021).

- Live Inputs: Exhibited detailed textural variations, including skin pores and wrinkles.
- Spoofed Inputs: Displayed smoother gradients and lacked natural edge discontinuities, making them easier to classify.

C. Scalability and Real-Time Performance

The system was tested for scalability and response times, processing inputs at an average of 20 frames per second on standard GPUs. This ensures its viability for real-time applications, such as integrating into biometric authentication systems.

D. Limitations

Certain challenges were observed:

- Partial Occlusions: Inputs with obstructed faces occasionally led to misclassification.
- Low-Resolution Inputs: Images below 100x100 pixels posed difficulties in extracting meaningful features.

These limitations will be addressed in future iterations to enhance model robustness.

IX. CONCLUSION

The proposed face presentation attack detection system leverages the principles of supervised learning and gradient- based analysis to identify live and spoofed inputs effectively. By combining texture extraction, edge irregularity analysis, and robust training techniques, the system achieves high accuracy and real-world applicability.

The research highlights the importance of addressing security vulnerabilities in biometric systems, especially for online authentication workflows. The methodology and results validate the feasibility of deploying this system in industries such as banking, e-commerce, and secure access control.

FUTURE SCOPE

To enhance the system's security and applicability, future research should focus on integrating multi-factor authentication (MFA) by combining face liveness detection with other biometric modalities like fingerprint or voice recognition. Additionally, incorporating 3D face analysis using depth- sensing cameras or infrared sensors will help detect sophisticated spoofing attempts, such as high-quality 3D masks (mays alshaikhli, 2021), by distinguishing real facial contours from artificial ones. Optimizing the model for edge devices through techniques like model pruning and quantization will enable efficient performance on low-power devices, making real-time applications more feasible (anjith geoge, 2019).

Expanding the dataset to include advanced spoofing attempts, such as high-resolution 3D masks and digitally manipulated videos, will improve the model's generalization capabilities. Collaborations with industry partners can provide access to larger, anonymized datasets for comprehensive testing. Moreover, real-world deployment and feedback from diverse demographic groups and environmental conditions will ensure inclusivity and fairness, refining the system for broader adoption and resilience against evolving threats.

REFERENCES

[1] Anjith Geoge, Z. M. (2019). Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network. IEEE Xplore, 15(10.1109/TlFS.2019.2916652), 42-55.

[2] Chaitanya Nagpal, S. R. (2019). A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing. IEEE Xplore, 1(10.1109/IJCNN.2019.8852422), 14-19.

[3] Junwei Zhou, K. S. (2020). Face Anti-Spoofing Based on Dynamic Color Texture Analysis Using Local Directional Number Pattern. IEEE Xplore(10.1109/1CPR48806.2021.9412323), 10-15.

[4] Liu, L. S. (2018). Face Liveness Detection Based on Joint Analysis of RGB and Near-Infrared Image of Faces. Symposium on Electronic Imaging, 30(10.2352/ISSN.2470-1173.2018.10.IMAWM-373), 373 - 379.

[5] Mays Alshaikhli, o. e.-m. (2021). The Deep Learning Method for Image Face Anti-Spoofing Detection. IEEE Xplore, 9(10.1109/EUVlP50544.2021.9484023), 23-25.

[6] Mostaani, Z. a. (2019). Wide Multi Channel Presentation Attack (WMCA). IEEE Transactions on Information Forensics and Security, 1, 15-17.

[7] Weihua Liu, Y. P. (2022). Spatio-Temporal-Based Action Face Anti-Spoofing Detection via Fusing Dynamics and Texture Face Keypoints Cues. IEEE Transactions on Consumer Electronics, 70(10.1109/TCE.2024.3361480), 2401-2413.

[8] Zitong Yu, Y. Q. (2023). Deep Learning for Face Anti-Spoofing: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(5), 5607-5631.